PROTEOMIC ANALYSES OF EXCRETORY SECRETORY PROTEINS (ESP) AND MEMBRANE PROTEINS OF *Entamoeba histolytica* HM1:IMSS

JORIM ANAK UJANG

UNIVERSITI SAINS MALAYSIA

2018

PROTEOMIC ANALYSES OF EXCRETORY SECRETORY PROTEINS (ESP) AND MEMBRANE PROTEINS OF *Entamoeba histolytica* HM1:IMSS

by

JORIM ANAK UJANG

Thesis submitted in fulfilment of the requirements
for the degree of
Master of Science

September 2018

ACKNOWLEDGEMENT

I would like to express my gratitude to Dr Nurulhasanah Othman, my supervisor, and Prof Rahmah Noordin, my co-supervisor. They have poured out their best in moulding me and guiding me in this journey to be a researcher. Not forgetting to mention, my personal development too. For that, I am forever grateful to them.

This research was supported/partially supported by Mr Muhammad Hafiznur Bin Yunus from the Insitute for Research in Moleculalr Medicine (INFORMM, USM), Dr. Mohd Nazri Ismail and Dr-to-be Sebastian Kwan from the Analytical Biochemistry Research Centre (ABrC, USM) as well as Pn Asmahani Azira Abdu Sani from the Malaysia Genome Institute (MGI, NIBM). Not to forget, all the staff of INFORMM, USM who have supported every facet of this project.

I would like to acknowledge all my friends and my family in INFORMM and at home, Kuching, who have supported me in my ups and downs, and who have tolerated my silliness. Without your moral support, I would not have continued this journey. Especially to mommy and daddy, thank you for your tremendous support in your love and encouragement; and money of course.

Last but not least, a big shout out to God almighty! Thank You for the favour and thank You for all the experiences I gained in this journey.

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

Page

LIST OF SYMBOLS

°C	degree Celsius
µL	microlitre
µg	microgram
g	gram
g	earth's gravitational acceleration/relative centrifugal force
L	litre
M	molar
mg	milligram
mL	millilitre
mM	millimolar
nm	nanometer
β	beta

LIST OF ABBREVIATIONS

1D	1-dimensional
2DE	2-dimensional electrophoresis
ACN	Acetonitrile
ALA	Amoebic liver abscess
APS	Ammonium persulfate
CHCA	α-Cyano-4-hydroxycinnamic acid
CID	Collision induced dissociation
CLS	Cyst-like structure
CSA	Crude-soluble antigen
DTT	Dithiothreitol
ECD	Electron-capture dissociation
ES	Excretory-secretory
ESA	Excretory-secretory antigen
ESI	Electrospray ionisation
ETD	Electron-transfer dissociation
FA	Formic acid
HCl	Hydrochloric acid
IAA	Iodoacetamide
LC	Liquid chromatography
m/z	mass-to-charge ratio
MALDI	Matrix-assisted laser desorption ionisation
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
MW	Molecular weight
NaOH	Sodium hydroxide
PAGE	Polyacrylamide gel electrophoresis

PBS	Phosphate buffer saline
PCR	Polymerase chain reaction
pI	Isoelectric point
PMF	Peptide mass fingerprinting
PPDK	Pyruvate phosphate dikinase
RCDC	Reducing agent and detergent compatible
RT	Real time
SD	Standard deviation
SDS	Sodium dodecyl sulfate
TEMED	Tetramethylethylenediamine
TFA	Trifluoroacetic acid
TOF	Time-of-flight
TOF/TOF	Tandem time-of-flight

ANALISIS PROTEOMIK PROTEIN PERKUMUHAN PEREMBES (ESP) DAN PROTEIN MEMBRAN DARIPADA *Entamoeba histolytica* HM1:IMSS

ABSTRAK

Entamoeba histolytica ialah sejenis parasit protozoa yang menyebabkan amebiasis. Jangkitan parasit ini boleh menyebabkan disentri dan abses hepar yang mana jangkitan tersebut seringkali menyebabkan kematian jika tidak dirawat. Hingga kini, kefahaman patogenesis tentang jangkitan ameba adalah terhad. Oleh itu, dalam kajian ini, analisis proteomik dilakukan untuk mengenal pasti sub-proteom protein perkumuhan perembes dan protein membran dalam trofozoit *E. histolytica*. Dari 209 protein, sebanyak 209 protein perkumuhan perembes telah dikenalpasti. Senarai ini masing-masing terdiri daripada 191 dan 97 protein yang berjaya dikesan oleh LC-ESI-MS/MS dan LC-MALDI-TOF/TOF. Sebanyak 79 protein telah dikenalpasti oleh kedua-dua sistem spektrometri massa tersebut, manakala 112 dan 18 protein masing-masing dikesan secara eksklusif oleh LC-ESI-MS/MS dan LC-MALDI-TOF/TOF. Ramalan *in-silico* mendapati 8 dan 31 protein masing-masing digolongkan sebagai protein perembes klasik dan tidak klasik. Klasifikasi ontologi menunjukkan peratusan besar iaitu sebanyak 23% protein perkumuhan perembes yang tergolong sebagai oxidoreductase. Seterusnya, fasa kedua kajian ini melibatkan perbandingan tiga kaedah pengekstrakan membrane iaitu dua kit komersil (ProteoExtract® daripada Calbiochem dan ProteoPrep® daripada Sigma), dan kaedah konvensional. Hasil kajian menunjukkan bahawa kit ProteoExtract® dan kaedah konvensional telah mengekstrak hasil protein yang lebih banyak berbanding dengan kit ProteoPrep®. Gabungan data daripada LC-MALDI-TOF/TOF dan LC-ESI-MS/MS telah

mengenalpasti masing-masing protein sebanyak 490, 492, dan 587 daripada ekstrak membran yang menggunakan kit ProteoExtract®, ProteoPrep®, dan kaedah konvensional. Analisis *in-siliko* meramalkan protein membran sebanyak 109 (22%), 237 (48%) dan 182 (31%) dalam ekstrak yang menggunakan kit ProteoExtract®, ProteoPrep® dan kaedah konvensional. Tambahan pula, pengenalpastian protein sitosol dan membran membuktikan bahawa kit ProteoPrep® merupakan kaedah yang paling selektif dan sensitif bagi pengekstrakan protein membran. Sebagai kesimpulan, hasil kajian ini telah membuktikan penemuan masing-masing 39 dan 249 senarai protein perkumuhan perembes dan protein membran *E. histolytica.* Tambahan pula, kajian ini telah mengesahkan bahawa penggunaan dua jenis spektrometer massa boleh meningkatkan liputan proteom. Kajian ini juga telah menambahkan pemahaman tentang jenis protein yang dikumuh dan dirembes oleh *E. histolytica* dan juga protein yang berada di membran parasit tersebut. Protein yang dikenal pasti sangat berguna untuk kajian selanjutnya bagi memahami penyakit amebiasis dan peranan protein ini dalam interaksi perumah dan parasit.

PROTEOMIC ANALYSES OF EXCRETORY SECRETORY PROTEINS (ESP) AND MEMBRANE PROTEINS OF *Entamoeba histolytica* HM1:IMSS

ABSTRACT

Entamoeba histolytica is a protozoan parasite that causes amoebiasis. Infection of this parasite may lead to amoebic dysentery and amoebic liver abscess, which is fatal if left untreated. Until now, understanding of the pathogenesis of amoebiasis is limited. Hence, in this study, proteomic analyses were performed on the excretory-secretory (ES) and the membrane sub-proteomes of *E. histolytica* trophozoites. A total of 209 ES proteins were identified in which 191 and 97 proteins were detected by LC–ESI–MS/MS and LC–MALDI–TOF/TOF, respectively. Of the 209 proteins, 79 were identified by both mass-spectrometry systems, while 112 and 18 proteins were detected exclusively by LC–ESI–MS/MS and LC–MALDI–TOF/TOF respectively. Subsequently, the secretome prediction analyses were performed whereby 8 and 31 out of 209 total proteins were identified as classically and non-classically secreted proteins, respectively. Functional annotation classification showed that the largest ES protein class, which is 23%, is the oxidoreductase. The second part of this study involved the comparison of three membrane protein extraction methods: two commercial kits (ProteoExtract® from Calbiochem and ProteoPrep® from Sigma), and a conventional laboratory method. The results showed that the ProteoExtract® kit and the conventional method extracted higher protein yields compared to the ProteoPrep® kit. The combined data from LC-MALDI-TOF/TOF and LC-ESI-MS/MS identified 490, 492, and 587 proteins extracted using the ProteoExtract®, ProteoPrep®, and conventional methods,

respectively. *In-silico* analysis predicted 109 (22%), 237 (48%) and 182 (31%) membrane proteins from the ProteoExtract®, ProteoPrep® and conventional method extracts, respectively. Furthermore, the identification of the cytosolic and membrane protein fractions showed that the ProteoPrep® extraction kit was the most selective and specific for the extraction of the membrane proteins. In conclusion, the results revealed 39 and 249 *E. histolytica* ES and membrane proteins, respectively. Furthermore, this study confirmed that the use of two types of mass spectrometers enhances proteome coverage. The data generated has increased the understanding on the types of proteins that are excreted-secreted by *E. histolytica* and also the proteins that reside at the parasite's membrane. The identified proteins will be useful for further studies in understanding the pathogenesis of amoebiasis and the roles the proteins play in the host-parasite interactions.

CHAPTER 1 – INTRODUCTION

1.1 An overview

Amoebiasis was first reported as a deadly disease in 1873 by Hippocrates who examined a patient suffering from bloody dysentery (Tanyuksel and Petri, 2003). Two years later, *Entamoeba histolytica* trophozoite was identified by Fedor Aleksondrovich Lösch in a farmer who suffered from a fatal case of dysentery (Marshall et al., 1997). Further investigation by inoculating the stool of the patient into the rectum of a dog caused a similar manifestation (Marshall et al., 1997). A significant milestone was achieved with the characterisation of *E. histolytica* as the causative agent for amoebic colitis and amoebic liver abscess (ALA) in the 1890s by Sir William Olser and his colleagues (Tanyuksel and Petri, 2003). Subsequently, the identification of cyst as an infectious stage was confirmed by Walker and Sellards in 1913, and followed by the establishment of the *E. histolytica* life cycle by Dobell in 1925 (Tanyuksel and Petri, 2003).

Most patients infected with *E. histolytica* are asymptomatic or only suffered from mild diarrhoea (Hankenson et al., 2003). Meanwhile, only 10% of the patients presented classic amoebic symptoms such as stomach cramps and bloody diarrhoea (Farthing, 2006). *E. histolytica* was not immediately associated as the causative agent of amoebiasis because most amoebic infections cases were asymptomatic. However, subsequent studies found that the infectious and the non-infectious amoeba were not similar (Fotedar et al., 2007). Since then, *E. histolytica* was reclassified into two species namely the infectious species, *E. histolytica* and the non-infectious species, *E. dispar* (Fotedar et al., 2007).

In 1997, amoebiasis was ranked second as death-causing parasitic infection, after malaria (World Health Organization, 1997). Approximately 40, 000 to 100, 000 deaths occurred annually, which include 1.9% to 9% of amoebic colitis patients (Aristizábal et al., 1991). Death occurrence in amoebic liver abscess (ALA) cases have decreased to 1 – 3% due to the effective medical intervention. Nonetheless, the mortality rate caused by the late detection resulting in the sudden intraperitoneal rupture occurred in 2 – 7% of the patients (Stanley Jr, 2003).

The results of previous studies have contributed to the advancement on many aspects in the management of amoebiasis. This includes a better way of diagnosis whereby the detection of pathogenic *E. histolytica* could be accurately distinguished from the morphologically similar but non-pathogenic *E. dispar* (Fotedar et al., 2007). Although many attempts have been made to improve the management of amoebiasis, the disease remains prevalent in underdeveloped countries of warmer climate (Walsh and Ravdin, 1988). Furthermore, the combination of poor sanitation and bad water quality provides the optimum breeding ground for this parasite (Walsh and Ravdin, 1988).

Large amounts of information on *E. histolytica* genome were made available since it was sequenced in the year 2005 (Loftus et al., 2005). In tandem, the advent of proteomic technologies has allowed proteomic studies on amoebiasis to be conducted. Early studies focused on analysing the subcellular expression profiles of trophozoites under various conditions (Davis et al., 2006, Tolstrup et al., 2007, Perdomo et al., 2015). Tolstrup et al. (2007) used 2-dimensional gel electrophoresis (2-DE) followed by mass spectrometry (MS) on 400–1500 *E. histolytica* protein spots. Therein, 63 proteins were identified and found related to cytoskeleton, surface, metabolic, the ubiquitin-proteasome system and signalling associated proteins. The

application of differential protein expression analysis using 2-DE and subsequent MS analysis was also performed to compare the proteome of Rahman and HM1:IMSS strains of *E. histolytica*. The results showed six proteins were found differentially expressed between the two strains (Davis et al., 2006). In 2015, a study on the *E. histolytica* trophozoite ER and Golgi apparatus using LC-MS/MS identified over 1,500 proteins of which are involved as trafficking machinery and GTPases (Perdomo et al., 2015). Hence, with the advancement of proteomic technologies and a complete *E. histolytica* protein database, high-throughput studies on the proteome of *E. histolytica* can be conducted.

1.2 Problem statements and rationale of the study

The shift of analysing a single protein to larger sets of proteins such as the excretory-secretory (ES) proteins and membrane proteins is made possible with advances in the proteomics technologies. Currently, there is no proteome report on the ES proteins of *E. histolytica*. Other than the study of *E. histolytica* cell surface membrane proteins by Biller et al. (2014), no other study has been performed on the membrane proteome of *E. histolytica*. Thus, this study aimed to identify the ES and membrane proteomes of *E. histolytica*. The identification of these proteins and their functions could add to the knowledge in understanding amoebic pathogenesis.

1.3 Objectives of the study

This study was conducted with the following objectives:

1. To perform proteomic analysis using LC-MALDI-TOF/TOF and LC-ESI-MS/MS and functional classification *via* PANTHERDB for the identified *E. histolytica* excretory-secretory proteins.
2. To compare *E. histolytica* membrane protein extraction methods: two commercial kits namely ProteoExtract® (Calbiochem), ProteoPrep® (Sigma), and a conventional laboratory method.
3. To perform proteomic analysis using LC-MALDI-TOF/TOF and LC-ESI-MS/MS and functional classification *via* PANTHERDB for the identified *E. histolytica* membrane proteins.

1.4 Workflow

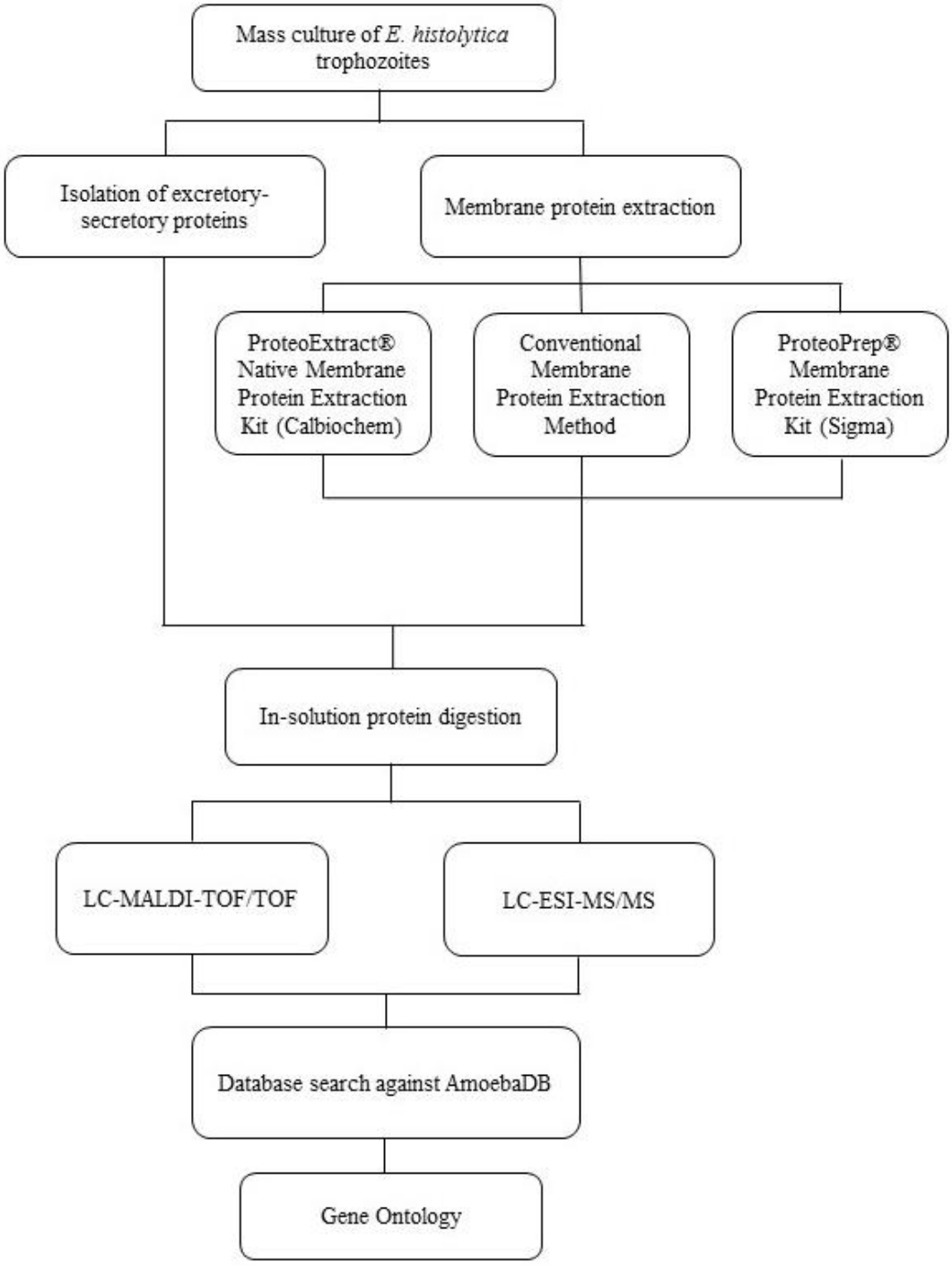

Figure 1.1 Flowchart of the study.

CHAPTER 2 – LITERATURE REVIEW

2.1 Biology of *E. histolytica*

2.1.1 Life cycle

E. histolytica exists in two distinct stages namely trophozoites and cysts. The simple life cycle begins with the consumption of the tainted fluid containing *E. histolytica* cyst (Hankenson et al., 2003). The cyst withstands harsh environment such as the gastric acid. Upon reaching a conducive environment such as the small intestine, a single cyst ex-cysts to form 8 trophozoites. These blood ingesting trophozoites then colonise the colon and cause dysentery. Trophozoites are unable to live in an unconducive environment outside the host or the host's gastric acids unless quadrinucleate cysts are formed again through a process known as encystation. Humans and primates are the only natural hosts for *E. histolytica* (Rivera et al., 2010, Stanley Jr, 2003). Figure 2.1 depicts the life cycle of *E. histolytica*.

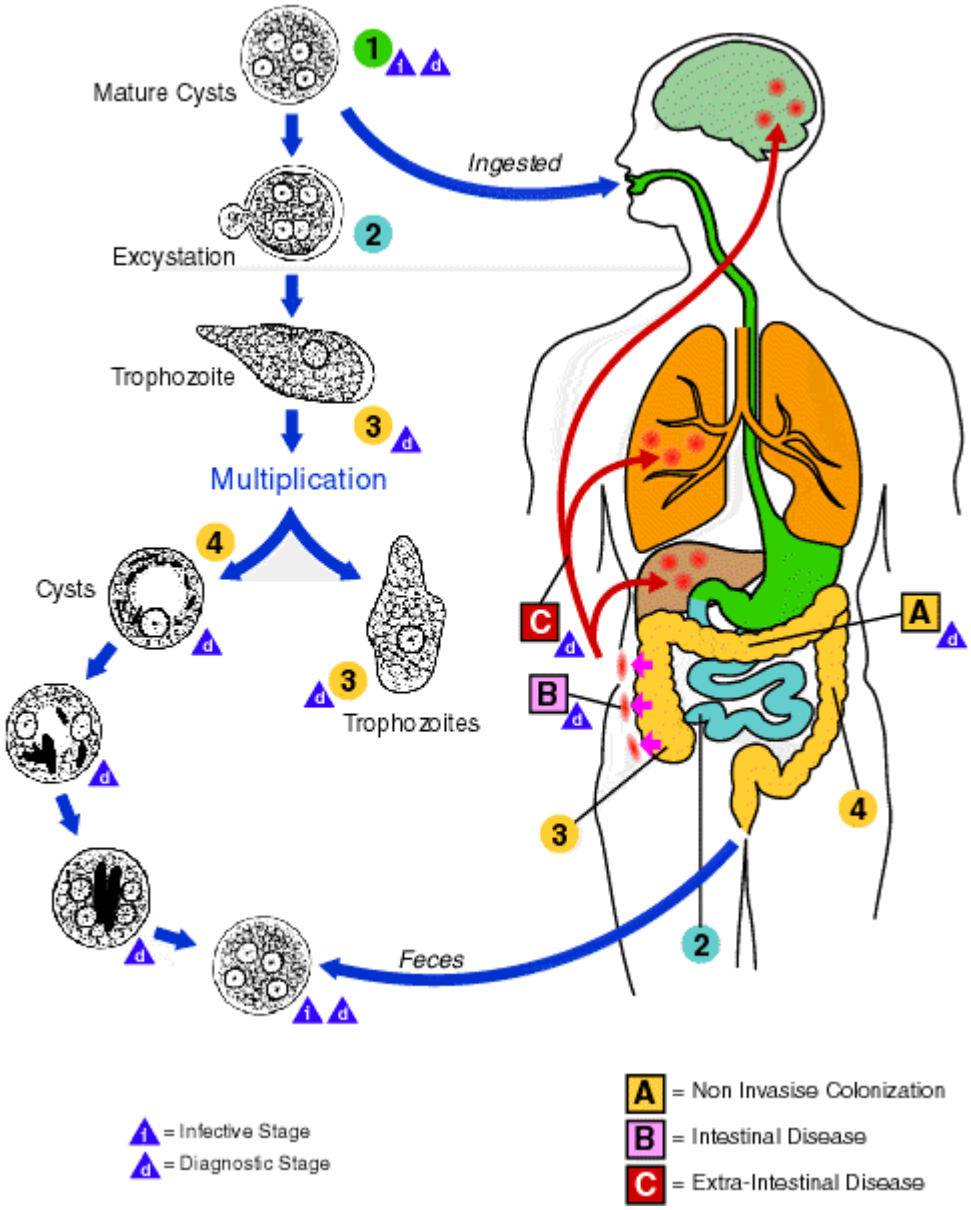

Image source:

https://www.pharmacology2000.com/Chemotherapy/Antiparasitic/AmebiasisLifeCycle.gif

Figure 2.1 The life cycle of *Entamoeba histolytica*.

2.1.2 Cell morphology

E. histolytica is a unicellular eukaryotic organism. The cyst form of *E. histolytica* is round in shape and is enclosed within a refractile wall which protects them from harsh conditions such as the stomach acid. It is responsible for the transmission of the disease. A mature cyst is 10-15 μm in size and consists of four nuclei. From a single quadrinucleate cyst, eight uninuclear trophozoites are formed through a process known as excystation. The trophozoite form of *E. histolytica* is 10-50 μm in size and consists of one nucleus. It is actively motile with finger-shaped pseudopodia and responsible for tissue invasion and damage. It is also responsible for causing tissue damage to the host (Stanley Jr, 2003).

In the trophozoite form, it contains a single nucleus and multiplies by binary fission. It is an endoparasite whereby it ingests nutrients from the host and can alter its shape for various purposes such as locomotion and evasion of the host immune responses (Espinosa-Cantellano et al., 1992, Markiewicz et al., 2011).

Other amoeba species such as *E. dispar* and *E. moshkovskii* share the same physical features with *E. histolytica*, thus causing difficulty in differentiating them from *E. histolytica* under the microscope (World Health Organization, 1997, Fotedar et al., 2007). The need to distinguish *E. histolytica* from other non-pathogenic *Entamoeba* species is important to avoid misdiagnosis and wrong treatment.

2.1.3 Transmission and occurrence

2.1.3(a) Susceptibility and risk factors

Typically, amoebiasis is acquired through the faecal-oral route, whereby food or water contaminated by the cyst form of *E. histolytica* is ingested by the host. Transmission can also occur through oral and anal sex as well as contaminated enema apparatus (Istre et al., 1982). According to Hankenson et al. (2003), the communicability of the disease is high as asymptomatic carriers can be a source of further infection. Furthermore, common household pest like flies and cockroaches can help spread the cyst form of *E. histolytica*. Adults and infants have similar chances of acquiring amoebiasis. However, according to a report by Hung, Chang & Ji (2012), men who have sex with men have a higher risk of being infected with *E. histolytica*.

Amoebiasis is still a major health problem especially among the aboriginals and communities living in the remote areas of Malaysia (Tengku and Norhayati, 2011). A study on the prevalence of *E. dispar*/*E. histolytica* among school children in the interior of Sabah showed that 83.8% of them had the infection (Mahsol et al., 2008). One of the leading factors causing high intestinal parasitic infections including *E. histolytica* in remote communities involves water sources (Duc et al., 2011). For instance, the transmission of parasites occurs in settings where a river contaminated with human and animal excretion is used interchangeably for agriculture, socio-economic and personal hygiene (Duc et al., 2011).

2.1.3(b) Epidemiology

Amoebiasis commonly occurs in populations living in tropical areas that lack proper sanitation. The disease is prevalent in developing countries such as Mexico, India, Africa and Malaysia (Centers for Disease Control and Prevention, 2015). In western Nepal, amoebic infection was ranked second after giardiasis (Mukhopadhyay et al., 2007). The prevalence of *E. histolytica* infection in the different regions of Brazil from the year 2001 to 2014 ranged between 6.8% and 46.3% (Silva et al., 2014). In Pakistan, the prevalence of *E. histolytica* was reported to be as high as 23.1%, whereby the most susceptible age group was found between 6 to 10 years old (Zeb et al., 2018).

In Malaysia, the prevalence of amoebiasis among the Orang Asli was found between 1% and 14% (Norhayati et al., 2003) (Figure 2.2). An outbreak of amoebiasis among the orang asli communities in the year 2004 reported 13.2% of 28 diarrhoea patients were infected by *E. histolytica* (Noor Azian et al., 2006). In 2005, 72.4% out of 58 liver abscess patients admitted to Hospital Universiti Sains Malaysia (HUSM) were found to be caused by *E. histolytica* (Zeehaida et al., 2008). Furthermore, between 2008 and 2009, 76.7% out of 30 liver abscess cases admitted to HUSM were also positive for *E. histolytica* DNA (Othman et al., 2010).

Despite being prevalent in tropical countries, human to human transmission can still occur regardless of climate and high sanitation standards. For example, in a temperate country such as Japan, mass *E. histolytica* infection at an institution for the mentally disabled in the Yamagata Prefecture of Japan reported 5 to 10% of people infected were symptomatic, while 90 to 95% of infected subjects were asymptomatic (Haghighi et al., 2003). In a report by Vreden et al. (2000), an amoebiasis outbreak in

the Netherlands demonstrated that *E. histolytica* can remain dormant for 13 years in their climate.

2.1.4 Disease, diagnosis and treatment

2.1.4(a) Symptoms

In most *E. histolytica* infections, symptoms are either not present or very mild (Stanley Jr, 2003). The majority of asymptomatic patients excrete cysts for a short period and are clear from the infection within 12 months of infection (van Hal et al., 2007). Only a small percentage of people infected with *E. histolytica* develops clinical symptoms. Patients with symptomatic amoebiasis often suffer from amoebic colitis and amoebic liver abscess (ALA) (Stanley Jr, 2003).

Patients with amoebic colitis commonly present a history of persistent abdominal pain and diarrhoea with the presence of blood and mucus in the stool. As amoebiasis is often neglected, a study reports that common inappropriate symptomatic treatment using corticosteroid has led to toxic megacolon complication in about 0.5% of patients (Ackers et al., 1997). Furthermore, when left untreated, the resulting gut perforation, exhaustion, and extraintestinal amoebiasis will lead to death (Hankenson et al., 2003).

ALA is the most common extraintestinal manifestation of amoebiasis (van Hal et al., 2007). As mentioned by Zurauskas & McBride (2001), patients who develop ALA are usually presented within 5 months of exposure to the disease, with clinical symptoms such as fever, and right upper abdominal quadrant pain. The majority of ALA patients do not present amoebic colitis symptoms and also *E. histolytica* cysts and trophozoites are rarely found in their stools (Fotedar et al., 2007).

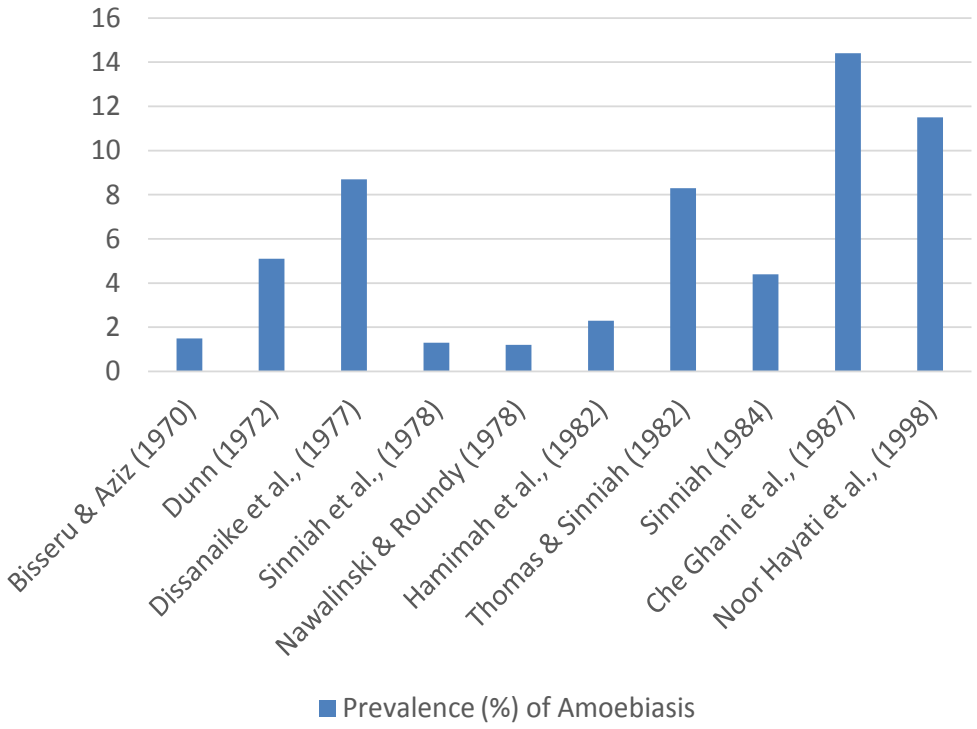

Figure 2.2 A compilation of studies by Norhayati et al. (2003) on the prevalence of amoebiasis among the orang asli communities in West Malaysia.

2.1.4(b) Diagnosis

The earliest diagnosis method of amoebiasis is the microscopic examination of stool samples whereby *E. histolytica* trophozoites can be seen containing red blood cells. However, this method is prone to cause misdiagnosis as other morphological similar strains, such as. *E. dispar* and *E. moshkovskii* are indistinguishable from *E. histolytica* under the microscope (Liang et al., 2009, Haque and Petri, 2006). Although microscopic method is routinely being used to diagnose amoebic colitis, it is not suitable to be performed for diagnosis of ALA cases. Despite low sensitivity of the microscopy, it is still being practised in many hospital laboratories.

Amoebic colitis patients can also be diagnosed by detecting small ulcers on colonic lesions obtained during the colonoscopic biopsy (Ohnishi et al., 2004). Furthermore, colonoscopy and subsequent sampling by means of culture swap are useful in patients with acute colitis and in cases when *E. histolytica* infection is suspected but failed to be detected in stool samples. However, these methods are time-consuming and sensitivity of the diagnosis is only 50% (Clark and Diamond, 2002). Antigen detection methods, eg. Entamoeba CELISA Path kit (Cellabs, Sydney, NSW) and the *E. histolytica* II kit (TechLab Inc, Blacksburg, Va, USA), are specific and can distinguish *E. histolytica* from *E. dispar*. The sensitivities and specificities of these various antigen detection kit ranges from 80% to 99% and from 86% to 98%, respectively (Haque et al., 1995, Gonin and Trudel, 2003, Furrows et al., 2004, Solaymani-Mohammadi et al., 2006). These tests are rapid and their interpretations are more definitive compared to the microscopic examination.

For the diagnosis of extraintestinal amoebiasis such as ALA, radiology imaging is used to detect the presence of an abscess in the liver. When the abscess is

present, further analyses such as culture, DNA detection, and/or antigen detection are performed. DNA and antigen detection-based methods performed on the abscess sample were reported to be highly sensitive (Fotedar et al., 2007, Paul et al., 2007, Tanyuksel and Petri, 2003). With serological methods, serum samples were used to detect antibodies against *E. histolytica* for the diagnosis of ALA. Commercial antibody detection assays made of native *E. histolytica* trophozoite antigens are available (Lotter et al., 1992, Ning et al., 2013). However, this method is ineffective to distinguish recent infection from past infection as high background antibody titre may persist in a population of endemic areas (Pillai et al., 1999, Zengzhu et al., 1999, Zeehaida et al., 2008, Mohamed et al., 2009).

Molecular diagnostic tests using polymerase chain reaction (PCR) to amplify *E. histolytica* DNA from the extracted faecal and pus of ALA patients are shown to be highly sensitive and specific (Gonin and Trudel, 2003, Solaymani-Mohammadi et al., 2006). Furthermore, the application of real-time PCR (RT-PCR) has significantly shortened detection time by simultaneous monitoring of the amplification process (Othman et al., 2010). The advantages of RT-PCR are the ability to detect a low number of parasite and the reliability in differentiating non-pathogenic *Entamoeba* species from *E. histolytica* (van Hal et al., 2007). However, these methods require skilled personnel and the high cost of reagents and equipment.

2.1.4(c) Treatments

Treatment for amoebiasis includes the oral administration of metronidazole or diiodohydroxyquin and in conjunction with a luminal agent such as iodoquinol (Hankenson et al., 2003). For patients with invasive amoebiasis, surgical drainage

may be unnecessary to treat ALA, as drug therapy alone is efficient (Akgun et al., 1999). However, aspiration of the abscess was shown to be beneficial in patients with large abscesses (Weinke et al., 2002). Meanwhile, asymptomatic carriers should be treated with a luminal agent to reduce the spread of disease and the risk of developing symptomatic infection (Stanley Jr, 2003).

Current drug therapies have been shown to cause several side effects. According to Petri Jr & Singh (1999), effective luminal agents such as diloxanide furoate and paromomycin caused frequent gastrointestinal disturbances and rare double vision, and symptoms related to ototoxicity and nephrotoxicity. Furthermore, other drugs used to treat amoebiasis such as metronidazole may cause unpleasant side effects, such as metallic taste, nausea and headache. Although it is uncommon, metronidazole can also cause neurological side effects, such as vertigo and encephalitis. Hence, treatments are discontinued whenever harmful side effects are present.

2.1.5 Pathogenesis

According to Lejeune, Rybicka, & Chadee (2009), *E. histolytica* trophozoites can maintain a commensal relationship with the host. This is generally observed in the infected individuals. When triggered, the unharmful relationship can turn destructive, beginning from the destruction of the intestinal wall, to the extent of invading surrounding soft organs such as the liver. Several works have been performed to distinguish the virulent and attenuated strains as well as to uncover the stress-inducing components from the tissue environment (Faust and Guillen, 2012). However, the conversion of the parasite from a commensal state to a destructive state

needs to be further elucidated. The current knowledge of the pathogenesis of intestinal amoebiasis is shown in Figure 2.3.

This host tissue lysing parasite has phagocytic, proteolytic, and cytolytic capabilities. Gal-lectin, cysteine proteinase and amoebapore are the three proteins known as the main culprit in the pathogenesis of amoebiasis. The invasion of the intestinal mucosa by the degradation of the mucin layer is thought to be the first strategy to disrupt the mucus gel (Moncada et al., 2005). In this strategy, the cysteine proteases secreted by *E. histolytica* effectively degraded the cysteine-rich domains of the MUC2 polymer of the mucus gel. Hence, it permits the parasite to come into contact with the epithelial surface (Lidell et al., 2006).

Then, the trophozoites attach to the tissue surface through its surface protein, namely Gal/GalNAc lectin. This leads to the cytolysis of the host cell (Tavares et al., 2005). A more recent study using an ex-vivo human intestinal model to study *E. histolytica* pathogenesis found that impairing the Gal/GalNAc lectin did not inhibit the parasite's attachment ability (Bansal et al., 2009). This study suggests that other molecules may also be involved in the adherence process. Subsequently, the destruction of the villin and microvilli occur as the trophozoites continue to secrete cysteine proteases.

The prevention of *E. histolytica* from invading the sub-epithelium region involves the production of nitric oxide (NO) and reactive oxygen species (ROS) in the activated macrophages. However, the parasite is able to defend itself from the attack by neutralising both NO and ROS with peroxiredoxin (Choi et al., 2005).

Various proteins are involved in progressing intestinal amoebiasis to extraintestinal amoebiasis. For example, the amoebapores play a crucial role in the

establishment of amoebic liver abscess (ALA) (Zhang et al., 2004). In addition, according to Santi-Rocca et al. (2008), the parasite upregulates the expression of lysine and glutamic acid-rich protein (KERP1) during ALA formation. The roles of these proteins are postulated in the protection of this parasite from an acute immune response during the development of ALA.

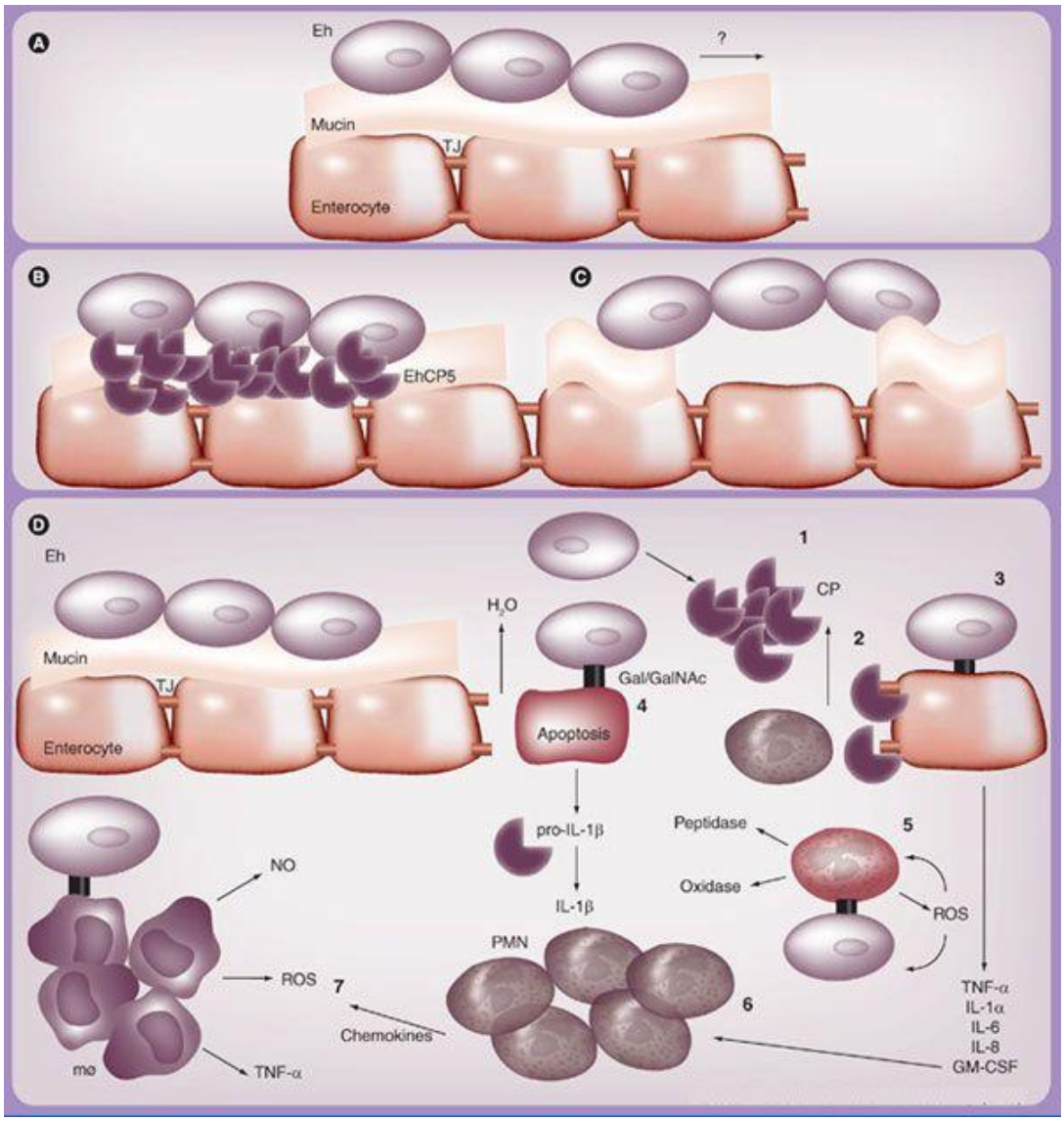

Figure 2.3 Pathogenesis of intestinal amoebiasis (Lejeune et al., 2009)

2.2 The proteome of *E. histolytica*

The term proteome was coined by Marc Wilkins in 1994 and it was defined as an entire complement of proteins expressed by the genome, cell, tissue or the entire organism (Wilkins et al., 2013). The overall aim of proteome studies is to perform large-scale discovery of the proteome by analysing many proteins at the same time. Large-scale proteomics approach, also known as a bottom-up strategy, has been proven to be an indispensable tool towards understanding the parasite pathogenesis (Veras and Bezerra de Menezes, 2016, Bertin et al., 2016).

For pathogens such as *E. histolytica*, its proteome is important for the pathogenicity of the disease and the cell viability. In amoebiasis, *E. histolytica* evades the immune system using surface receptor capping on the uropods. In this event, the targeted host immune components on the parasite's surface are translocated and shed at the uropod (Avila-Calderón et al., Espinosa-Cantellano et al., 1992, Markiewicz et al., 2011). The shedding of uropods from the parasite suggests that the isolated fraction contains various molecules at the plasma membrane. The uropod shedding also suggests its involvement in the excretory mechanism of the trophozoites. In a study to understand the mechanism of uropod formation, the identification of the uropod proteome showed several numbers of multiple drug resistance proteins, ATPases, GTPases, and cysteine proteases (Markiewicz et al., 2011).

The *E. histolytica* proteome plays an important role in the encystation and excystation in response to the environment. Proteome analysis of the total *E. histolytica* protein identified 1029 proteins from the trophozoite form, 550 proteins from the cyst-like structure (CLS), and 411 proteins from the cysts, with only 74

proteins found to be common across all the three forms (Luna-Nácar et al., 2016). This study suggests that CLS may be an intermediate survival strategy of trophozoites towards stressful condition, of which the process enables the parasite to form a chitin-like resistant cover containing Jacob protein as a shelter.

During host infection, *E. histolytica* is exposed to reactive oxygen species that are released by the host's immune cells at the site of the infection. Shahi et al. (2016) identified 154 oxidising proteins, in which these proteins were involved in transport, catalysis, antioxidant activity, and maintaining the parasite's cytoskeleton. They also reported the involvement of arginase in the protection of the parasite against oxidative stress that was induced by the host. These results emphasise the contribution of oxidative stress by the host cells to the pathogenesis of *E. histolytica*.

Proteome analysis aimed to elucidate the migration-related proteins have identified EhPC4 (positive coactivator 4) to be responsible for the underlying mechanisms of *E. histolytica* trophozoites migration (de la Cruz et al., 2014). They have identified 16 differentially expressed proteins, of which four up-regulated proteins were involved in cytoskeleton organisation and cell migration. They observed that the overexpression of EhPC4 induced a significant increase in the trophozoite migration and the destruction of human SW480 colon cells. Hence, these proteins play an important role in the virulence of *E. histolytica*.

Many *E. histolytica* proteins play important roles in the parasite's pathogenesis, either by direct involvement or by important intracellular process. The PI3K family of intracellular signalling enzymes play a role in the early stages of phagosome formation (Powell et al., 2006, Nakada-Tsukui et al., 2009). Further proteome analyses on the phagosome have identified many GTPase families

(Rodríguez et al., 2000, Okada et al., 2006, Hernandes-Alejandro et al., 2013). These studies indicated that many protein members of the *E. histolytica* proteome is required for amoebic trogocytosis and phagocytosis to occur. In addition, the GTPase families include the Rab proteins, such as EhRab7A, EhRabA and EhRabB are localised at the phagocytic cup and may be part of the *E. histolytica* secreted proteins (Ralston, 2015).

2.2.1 Excretory-secretory proteins

During infection, *E. histolytica* trophozoites release excretory-secretory (ES) proteins, which are also known as excretory-secretory antigens (ESA). ES proteins are involved in the invasion of trophozoites into the colonic mucosa by degrading the glycoside substrates and proteins of the host tissues (Keene et al., 1986, Scholze and Werries, 1986, Reed et al., 1993, Moncada et al., 2005). Antibodies against ES proteins have been detected in the sera of both symptomatic and asymptomatic patients who have contracted amoebiasis (Pal et al., 1996).

The use of ES proteins as potential targets for diagnosis, treatment, and vaccine development for amoebiasis has been explored in previous studies (Quach et al., 2014, Wong et al., 2011, Saidin et al., 2014, Debnath et al., 2012). In diagnostics, the *E. histolytica* Gal/Gal-NAc lectin antigen is utilised in commercial antigen detection tests, i.e., the TechLab *E. histolytica* II ELISA (TechLab Inc). Furthermore, Gal/Gal-NAc lectin also showed potential as a vaccine candidate against *E. histolytica* (Quach et al., 2014). Another study on ES proteins showed the diagnostic potential of pyruvate phosphate dikinase (PPDK), and its recombinant form was used to develop a lateral flow dipstick test (Wong et al., 2011, Saidin et al., 2014). In

addition, auronofin was identified as an effective drug which targeted *E. histolytica* thioredoxin reductase (Debnath et al., 2012).

Proteome analysis on the ES proteins of *Trypanosome* sp. has uncovered a range of proteins which include unfolding and degradation classes of proteins, such as serine, cysteine proteases, and metallopeptidases (Nten et al., 2009). These proteases play a part in the physiological and pathological functions that favour the invasion of the parasite, growth in hostile host conditions, evasion of components of the host immune defence, and hydrolysis of host proteins.

E. histolytica secretome comprises a spectrum of proteins that may be needed for every facet of the parasite's life cycle including cell modulation that is due to the environmental adaptation and the evasion of host's immune responses. Hence, proteome studies of ES proteins may open paths to initiate novel strategies for the management and prevention of amoebiasis (Ahn et al., 2018).

2.2.2 Membrane proteins

Membrane proteins are important in many processes ranging from basic cellular process to self-defence and disease-causing processes (Santoni et al., 2000). Many of them are also potentially good drug targets, with an estimation of more than half of all drugs that have been developed targets the membrane proteins (Klabunde and Hessler, 2002). During *E. histolytica* infection, the parasite's membrane proteins are used for tissue invasion, as well as the establishment of intra and extra-intestinal infections. Biller et al. (2014) reported that the surface proteome of *E. histolytica*

consisted of 693 proteins, whereby 87% of the identified proteins were estimated to be localised on the membrane surface.

Perdomo et al. (2015) identified more than 1500 *E. histolytica* endomembrane proteins. The top two classes of proteins were involved in trafficking machinery and GTPases proteins with 152 and 131 proteins, respectively. The analysis revealed a high abundance of proteins that were involved in the intracellular trafficking mechanism. The most abundant protein was calreticulin, which resided at the endoplasmic reticulum and functioned as a calcium-buffer and a chaperone. It was localised at the plasma membrane and it was involved in the host cell interaction and the formation of phagocytic cups (Short et al., 2005).

The membrane proteome is a landfill for the exploitation of biomarkers. In a proteome analysis by Che et al. (2011), over two thousand *Toxoplasma gondii* membrane proteins were identified. Over 40% of the identified membrane proteins were hypothetical. Furthermore, many of the membrane proteins identified were unique to *T. gondii*. Hence, the study provided a set of proteins that are suitable for further experimental investigation.

The importance of identifying and studying membrane proteins is highlighted by the fact that they account for 70–80% of all drug targets. In addition, it is estimated that the majority of future drug targets are the membrane proteins (Hopkins et al., 2006, Overington et al., 2006). Therefore, the study of membrane protein in *E. histolytica* may even precede over ES proteins in search of potential novel biomarkers for drug targets. However, the detection of membrane proteins by standard proteomic methods is challenging due to the low abundance of membrane proteins relative to the total cell lysates, and their hydrophobic characteristics

(Santoni et al., 2000, Wallin and Heijne, 1998). Therefore, it is important to establish a protocol for isolation of membrane proteins prior to mass spectrometry analysis.

2.3 Tools for proteome discovery

2.3.1 Sample preparation

In-gel and in-solution digestions are the two common approaches in a bottom-up proteomic sample preparation (Figure 2.4). However, in the context of a complex protein sample, in-solution digestion requires a post peptide separation while in-gel digestion is already the result of pre-protein separation (Gundry et al., 2010).

In-gel digestion followed by a mass spectrometry analysis are widely used techniques to identify proteins (Lasonder et al., 2002, Nten et al., 2009, Pomastowski and Buszewski, 2014). Before protein digestion, separation of the protein is performed using sodium dodecyl sulfate in a polyacrylamide gel (SDS-PAGE). Then, individual protein band or spot can be cored out to proceed with in-gel digestion. Also, depending on the complexity of the sample, several strategies are required for an efficient mass spectrometry analysis. This includes the need to consider the application of either one or two-dimensional gel electrophoresis (2-DE) (Pomastowski and Buszewski, 2014).

In-solution digestion followed by mass spectrometry is one of the simplest and commonly used techniques (de Souza et al., 2006, Biller et al., 2014, Perdomo et al., 2015). This technique involves denaturing, reducing, alkylating, and digesting the protein sample in the liquid phase. The fractionation is usually performed after the digestion. Nonetheless, separation can also be performed prior to digestion using

different forms of chromatography tools, including, reverse-phase, strong and weak ion exchange, as well as size exclusion chromatography (Mostovenko et al., 2013).

The in-gel digestion has several advantages over the in-solution digestion. In this method, sample complexity can be reduced without using liquid chromatography (LC). On the other hand, the in-solution digestion method requires fractionation using LC after the sample digestion step. Hence, as increased sample complexity and the efficiency of protein identification require longer mass spectrometry time, cost-benefit may favour gel-based mass spectrometry compared to in-solution digestion (Rabilloud and Lelong, 2011). The in-solution digestion has its own advantages such that it is simple and straightforward to perform. Furthermore, the sample recovery of in-gel digestion is estimated to be 70 – 80% of the in-solution efficiency (Shevchenko et al., 2006, Gundry et al., 2010). In addition, the protein sample concentration and amount for in-solution digestion are fixed and hence the protein quantity can be controlled. However, for in-gel digestion, the amount of proteins digested from the gel is difficult to ascertain, though the amount of the initial protein load can be controlled (Zhou et al., 2005).